

ПРОГНОЗУВАННЯ РІВНЯ УЧАСТІ ЕЛЕКТОРАТУ В ГОЛОСУВАННІ ЯК ФАКТОР ТОЧНОСТІ ПРОГНОЗІВ РЕЗУЛЬТАТІВ ВИБОРІВ

У статті проаналізовано, як точність прогнозування рівня участі електорату у виборах впливає на загальну точність прогнозування результатів виборів. Розглянуто методи прогнозування «ймовірного електорату», зроблено огляд змінних, які можуть бути використані для розв'язання цього завдання. Наведено результати авторської побудови інструментів для прогнозування «ймовірного електорату» на даних українських досліджень.

Ключові слова: прогнозування результатів виборів, прогнозування явки, «ймовірний електорат».

Прогнозування рівня участі електорату у виборах («явки») є важливим і самостійним дослідницьким завданням соціологів. Але помилки в прогнозах явки також впливають і на точність прогнозування результатів виборів, особливо, коли дослідник помиляється у визначенні того, хто саме з респондентів передвиборчого опитування візьме участь у виборах. Завдання прогнозування явки можна розділити на дві складові: 1) прогнозування загальної виборчої активності; 2) прогнозування того, які саме групи виборців візьмуть участь у голосуванні. Ця стаття присвячена розв'язанню проблеми прогнозування того, які саме групи виборців візьмуть участь у голосуванні. За результатами аналізу факторів, що впливають на точність прогнозування результатів виборів у США, Креспі вказує, що точність прогнозу рівня участі електорату у виборах має одну з найбільших кореляцій із точністю прогнозу результатів голосування [1, с. 37]. Двома головними причинами помилок прогнозування як загального рівня участі у виборах, так і того, які саме групи візьмуть участь у голосуванні є: 1) низька прогностична валідність запитання про наміри брати участь у виборах (далі будемо називати його *інтенційним запитанням*); 2) нерепрезентативність вибірок передвиборчих досліджень, яка виникає внаслідок наявності зміщення, що викликане питаннями без відповідей. Українські соціологи найчастіше застосовують тільки один індикатор того, чи візьме респондент участь у виборах – відповідь на запитання: «Чи збираєтеся Ви взяти участь у виборах?». Пряме використання відповідей на це запитання не дає достатньої можливості прогнозувати як загальну явку, так і ймовірність кожного з респондентів проголосувати. Іntenційне запитання не має достатньої прогностичної валідності. Фрідман і Голдштейн продемонстрували, що ті, хто не збирався йти на вибори, справді не голосують, тоді як поміж тих, хто за-

явив про свою участь у виборах, далеко не всі проголосували (див. таблицю 1) [2, с. 578].

Таблиця 1. Зіставлення намірів щодо участі у виборах та реальної поведінки на даних National Election Study (в клітинах таблиці подані відсотки від усіх виборців, у дужках – по рядку)

| Поведінка \ Набір | Взяли участь у виборах | Не взяли участь у виборах |
|---------------------------|------------------------|---------------------------|
| Не брати участь у виборах | 1,3 % (6,5 %) | 18,5 % (93,5 %) |
| Взяли участь у виборах | 56,3 % (70,4 %) | 23,9 % (29,6 %) |

Але дані інших досліджень, зокрема й українських¹, демонструють, що ситуація складніша – навіть поміж тих, хто відмовився голосувати, значна частина респондентів бере участь у виборах (див. таблиці 2, 3). Іntenційне запитання недостатньо прогнозує як поведінку тих, хто каже, що буде голосувати, так і тих, хто не має наміру брати участь у виборах (ті самі тенденції зафіксовано нами на даних англійських досліджень).

Таблиця 2. Зіставлення намірів щодо участі у виборах та реальної поведінки на даних українського панельного дослідження, перший тур виборів Президента 2004 року (в клітинах таблиці подані відсотки від усіх виборців, у дужках – по рядку)

| Поведінка \ Набір | Взяли участь у виборах | Не взяли участь у виборах |
|--------------------------------------|------------------------|---------------------------|
| Точно не буду брати участь у виборах | 1,6% (38,4%) | 2,5% (60,3%) |
| Радше не буду брати участь у виборах | 1,1% (86,4%) | 0,2% (13,6%) |
| Важко сказати | 3,6% (70,8%) | 1,4% (27%) |
| Радше буду брати участь у виборах | 10,9% (84,3%) | 2,0% (15,3%) |
| Точно буду брати участь у виборах | 71,1% (92,8%) | 5,3% (7%) |

¹ Аналізуються результати панельного дослідження до і після голосування в першому турі виборів Президента України в 2004 році, що було проведено Київським міжнародним інститутом соціології спільно з Київським центром політичних досліджень і конфліктології.

Зауважимо, що абсолютна кількість тих респондентів, які в передвиборчому опитуванні відповіли, що точно не будуть брати участь у виборах та радше не будуть брати участь у виборах, невелика – 73 та 22 людини

Таблиця 3. Зіставлення намірів щодо участі у виборах та реальної поведінки на даних канадського панельного дослідження, 2006 рік ¹

| Поведінка Намір | Взяли участь у виборах | Не взяли участь у виборах |
|----------------------------------|------------------------|---------------------------|
| Точно не буду голосувати | 1,2 % (21,4 %) | 23,6 % (78,6 %) |
| Малоймовірно, що буду голосувати | 2,0 % (42,0 %) | 18,2 % (58,0 %) |
| Імовірно, що буду голосувати | 18,0 % (77,8 %) | 33,3 % (22,1 %) |
| Точно буду голосувати | 75,5 % (96,1 %) | 20,1 % (3,9 %) |

Проаналізувавши дані панельних електоральних досліджень, ми можемо стверджувати, що, на відміну від результатів, отриманих у роботі [2], інтенційне запитання неефективно передбачає поведінку навіть тих респондентів, які не збираються голосувати. Під час першого туру виборів Президента України 38 % тих, хто точно не збирався голосувати, усе ж таки взяв участь у виборах і 86 % тих, хто в передвиборчому опитуванні дав відповідь «скоріше не буду брати участь у виборах», теж проголосували (хоча, звісно, потрібно зважати на невелику абсолютну кількість опитаних). На виборах 2006 р. у Канаді проголосували 21 % тих, хто в першій (передвиборній) хвилі дослідження дав відповідь «точно не буду голосувати» і 42 % тих, хто відповів, що малоймовірно проголосує (і тут кількість опитаних уже набагато більша: 203 респонденти, які точно не збирались голосувати та 222 тих, хто відповів, що малоймовірно проголосує). На виборах в Англії спостерігалася така сама тенденція – з-поміж трьох груп із найменшою задекларованою ймовірністю проголосувати взяли участь у виборах відповідно 26 %, 30 % і 47 % ². Як можна побачити з наведених даних, інтенційне запитання також неефективно передбачає поведінку тих, хто збирається голосувати. Хоча, звісно, імовірність респондента справді взяти участь у голосуванні сильно пов'язана з його впевненістю, але не повністю її детермінує (див таблиці 1, 2, 3). Водночас, треба брати до уваги, що поствиборне панельне опитування – не дуже

надійний інструмент для вимірювання реальної явки на вибори через те, що в ньому так само діє механізм надання невірної інформації про участь у виборах (тобто, частина респондентів каже, що взяла участь у виборах, хоча вони не голосували) як і у передвиборних опитуваннях [3, с. 25]. У цьому можна переконатися, якщо порівняти частку тих, хто збирався голосувати, але не зробив цього, на даних [2] (у цій роботі валідувався факт голосування за даними офіційної статистики) та отриманих нами в ході аналізу українських, канадських та англійських поствиборних досліджень. Як видно, якщо використується валідація за офіційними даними, частка тих, хто збирався голосувати, і зробив це складає 56 %, а якщо аналізуються дані панельних досліджень – 84–93 %. Ми можемо також порівняти реальну явку на українських виборах та відповіді респондентів у передвиборних опитуваннях (див. табл. 4).

Таблиця 4. Розподіл відповідей на інтенційне питання порівняно з реальною явкою за даними Київського міжнародного інституту соціології.

| | 2002 рік | 2004 рік | 2007 рік |
|---|------------|------------|------------|
| Звичайно (безумовно, точно і т.д.) НІ | 7% | 5% | 10% |
| Радше (напевне і т.д.) НІ | 1% | 1% | 5% |
| Не знаю (ще не вирішив(ла), важко сказати і т.п.) | 6% | 5% | 6% |
| Радше (напевне і т.п.) ТАК | 13% | 13% | 17% |
| Звичайно (безумовно, без сумніву і т.п.) ТАК | 73% | 76% | 60% |
| Сума ствердних відповідей | 86% | 89% | 77% |
| Рівень участі у виборах | 69% | 74% | 58% |

Очевидно, якщо вважати, що у виборах візьмуть участь усі респонденти, котрі не відмовилися від цього, то ми значно переоцінимо реальний рівень участі. Якщо припустити, що проголосують тільки ті, хто дав ствердну відповідь – також. До того ж, немає підстав вважати, і це демонструють проаналізовані нами вище дані панельних опитувань, що ті, хто обрав альтернативу «безумовно так», взяли участь у виборах, тоді як всі респонденти інших категорій цього не зробили. Як демонструє проведений нами аналіз, сукупність реальних виборців є «сумішшю» із респондентів, які дали різні відповіді на інтенційне запитання. Використання тільки інтенційного запитання не є надійним способом визначення ймовірності респондента взяти участь у виборах, хоча, як вказують багато авторів і демонструє наш аналіз, саме наміри респондента мають найбільшу кореляцію з фактом голосування [4, с. 10]. Чому використання тільки інтенційного запитання, і, як наслідок, помилки в передбаченні того, хто саме проголо-

¹ Доступ до даних можна здійснити за покликанням <http://ces-ec.mcgill.ca/surveys.html#2006>

² Але, треба зазначити, що виборці, які мали намір не голосувати, але все ж таки пішли на вибори, складають незначну частку серед реальних виборців – близько 3 % в Україні, 3,2 % в Канаді. Тому неврахування цих голосів, скоріше за все, не може серйозно вплинути на якість прогнозу результатів виборів.

сує, загрожують точності прогнозування результатів виборів? Наш аналіз поствиборного дослідження 2002 р., продемонстрував, що 24 % респондентів не брали участі у виборах. З їх числа 63 % на запитання про причину неучасті у виборах відповіли «збирався, але перешкодили деякі свої справи». Якщо перерахувати цю кількість стосовно до всіх виборців, то вона складе 14,5 %. Респондентам, які ствердно відповіли на питання про участь у виборах, ставили питання про те, за кого вони віддали свій голос. Тих респондентів, які збиралися піти голосувати, але не пішли, питали, за кого вони хотіли проголосувати. За критерієм χ^2 ці два розподіли статистично значуще (на рівні 0,01) відрізняються. Серед тих, хто збирався, але не проголосував, було менше прихильників БЮТ, СПУ та СДПУ(о), і більше тих, хто хотів проголосувати «проти всіх». Зауважимо, що, коли порівняти аналогічні розподіли за категоріями відповідей на інтенційне запитання про участь у виборах: «безумовно так», «радіше так» «ще не вирішив», то відмінностей у них не буде, тоді як між «справжніми виборцями» та тими, хто не взяв участі у виборах, вони є. Це є ще одним свідченням того, що підгрупа тих, хто не проголосував, формується з респондентів, які дали *різні відповіді* на інтенційне запитання. Ще одним свідченням того, що підвибірки респондентів, які виокремлюються як потенційні виборці на основі інтенційного запитання, не є точним відображенням сукупності реальних виборців, є порівняння їх із даними екзитполу. Так, наприклад, КМІС на виборах 2002 р. визначав тих, хто голосуватиме, як сукупність респондентів, які відповіли, що вони точно або напевно братимуть участь у виборах [5]. ФОМ приписував кожному респондентові ймовірність проголосувати, залежно від міри його впевненості в тому, що він піде на виборчу дільницю¹. Чи є виокремлені таким чином підвибірки виборців репрезентативними для тих, хто справді пішов голосувати? Відповідь на це запитання знову ж таки можна отримати, аналізуючи дані екзитполу, тобто відповіді реальних виборців – якщо в передвиборному дослідженні та екзитполі ми маємо справу з тією самою сукупністю, то принаймні їхні демографічні характеристики повинні збігатися. Наш аналіз показав, що підвибірка респондентів передвиборного опитування КМІСу, які були впевнені, що підуть голосувати або відповіли, що напевно голосуватимуть² статистично значуще відрізняється від сукупності респондентів екзитполу (тобто реальних виборців) за статтю та ос-

вітою³ (відмінності зафіксовані нами також і під час аналізу виборів 2006 р.). Таку саму ситуацію ми спостерігаємо, якщо застосувати підхід ФОМу до виокремлення тих, хто голосуватиме. Відзначимо, що на подібну проблему вказують, наприклад, Белтран і Вальдіва [6, с. 130], аналізуючи причини неточності прогнозування виборів у Мексиці. Очевидно, що це серйозна методологічна проблема – прогнози результатів виборів розраховуються на підвибірці респондентів, яка не повністю відповідає сукупності тих, хто реально бере участь у голосуванні. Попри те, що розподіл відповідей на запитання щодо намірів брати участь у виборах справді дає змогу оцінити реальну явку (пов'язаний із нею кореляційною залежністю), для успішного прогнозування виборів важливішим завданням, як нам видається, є надійне визначення ймовірності кожного з респондентів проголосувати.

Огляд робіт, присвячених формуванню «ймовірного електорату»

Як розв'язується завдання точнішого передбачення того, хто саме візьме участь у виборах? Існують такі способи формування «ймовірного електорату»:

1) розроблення індексу (наприклад, за шкалою Гуттмана або адитивного), який складається з відповідей на запитання, що пов'язані з ймовірністю взяти участь у виборах. До складу «ймовірного електорату» зараховують тільки виборців, кількість ствердних відповідей яких перевищує певне значення. Припускається, що саме ці респонденти візьмуть участь у виборах, а ті, у кого значення нижче порогового – не будуть голосувати.

3) Розроблення математичної моделі (найчастіше використовується рівняння логістичної регресії), яка, оцінюючи набір характеристик респондента відповідно до їх значень, приписує кожному респонденту ймовірність взяти участь у виборах. Потім відбувається зважування вибірки відповідно до розрахованих ймовірностей. Тобто, до «ймовірного електорату» зараховують усіх респондентів, але з вагами, пропорційними ймовірності проголосувати.

Побудова індексу ймовірності проголосувати

Основа цього методу – використання шкали Гуттмана, яка складається із запитань, пов'язаних із ймовірністю взяти участь у виборах. Для побудови шкали необхідно мати можливість проаналізувати, який відсоток виборців у групах з однаковим значенням індексу справді проголосували. Для розв'язання цього завдання можуть

¹ 0,95 – для респондентів, що засвідчили, що неодмінно голосуватимуть, 0,2 – для тих, хто напевно піде голосувати, 0,15 – для тих, хто ще не визначився тощо.

² Тобто ті респонденти, на яких розраховувався прогноз результатів виборів.

³ Також є значуща відмінність за віком та регіоном проживання.

Таблиця 5. Змінні, які використовуються під час розрахунку ймовірності того, що респондент візьме участь у виборах [2; 7; 8; 9]

| | Траджетт і Такер (Traugott and Tucker) | Перрі (Perry) | Фрідман і Голдштейн (Freed- man and Goldstein) | Петрович (Petrocik) |
|--|--|------------------|--|------------------------|
| Який підхід використовувався – індекс чи логістична регресія | Індекс і логістична регресія | Індекс | Логістична регресія | Логістична регресія |
| Змінні-ставлення | | | | |
| Рівень зацікавленості респондента в поточній виборчій кампанії | + | + | - | - |
| Інтенсивність намірів респондента брати участь у виборах | - | + | - | + |
| Рівень зацікавленості в політиці | - | + | - | - |
| Рівень впевненості у своєму виборі (кандидата) | - | + | - | - |
| Сильна партійна ідентифікація | - | - | + | - |
| Рівень знань про політику | - | - | + | - |
| Змінні, які вимірюють поведінку | | | | |
| Чи є респондент зареєстрованим виборцем | + | + | - | - |
| Чи брав респондент участь у попередніх виборах | + | + | - | - |
| Питання про те, чи знає респондент, де знаходиться виборча дільниця | - | + | - | - |
| Питання про те, чи голосував уже респондент на виборчій дільниці, на території якої він зараз проживає | - | + | - | - |
| Як часто респондент голосував на виборах, які відбувалися раніше | - | + | - | - |
| Чи є респондент постійним відвідувачем церкви | - | - | + | + |
| Соціально-демографічні змінні | | | | |
| Чи одружений респондент | - | - | + | - |
| Рівень освіти | - | - | + | + |
| Вік | - | - | + | + |
| Регіон проживання | - | - | + | - |
| Раса | - | - | + | - |
| Чи є респондент власником житла | - | - | + | - |
| Чи живе респондент у теперішньому помешканні більше 2 років | - | - | + | - |

бути використані 1) офіційні дані про факт голосування конкретного виборця; 2) дані поствиборного панельного дослідження, яке дає змогу отримати цю інформацію від самого респондента. Цей метод почали активно розробляти в 60-ті роки в США, піонером у побудові подібних шкал був Рол Перрі, ще один варіант шкали запропонували Траджетт і Такер [7; 8]. І автори підходу, і його критики вказують на такі недоліки: індекс не є формальним методом передбачення рівня участі у виборах – від дослідника залежить, респондентів з якими значеннями індексу зараховувати до «ймовірного електорату»; відкидаючи виборців із низькою ймовірністю проголосувати, дослідник може втратити важливу інформацію, адже й поміж них є ті, хто насправді бере участь у виборах [7]. Як зазначає Петрович, методологія Траджетта-Такера виявляється не набагато кращою, ніж використання тільки інтенційного запитання [9, с. 643]. Причину він убагаває в тому, що запитання, відповіді на які використовуються для побудови показника

ймовірності проголосувати, мають тенденцію до зміщення в бік соціально-схвальних відповідей, як і власне інтенційне запитання. Водночас, використання такого індексу дає змогу досліднику аналізувати, як залежить прогнозний розподіл голосів від рівня явки.

Побудова рівняння логістичної регресії для розрахунку ймовірності взяти участь у виборах

Другий спосіб побудови «ймовірного електорату», порівняно з описаними вище є об'єктивнішим і менше залежить від зміщень у відповідях, що викликані сенситивним характером як інтенційного запитання, так і запитань, що застосовуються під час побудови індексу. Основою цього підходу є порівняння характеристик респондентів, які збирались взяти участь у виборах і зробили це, і тих, хто, усупереч своїм намірам, не проголосував. На основі цього порівняння будується рівняння логістичної регресії, за допо-

могою якого розраховується ймовірність взяти участь у виборах кожного респондента. При застосуванні регресії всіх респондентів зараховують до складу «ймовірного електорату» – розраховані за допомогою логістичної регресії значення ймовірності використовуються як ваги. Наскільки ефективними є запропоновані методи? Як указують американські дослідники, рівень точності запропонованих індексів та рівнянь складає від 73 до 78 % [10, с. 161].

Огляд змінних, які використовуються для оцінки ймовірності взяти участь у виборах

Усі змінні можна розділити на три класи: 1) змінні-ставлення (інтерес до політики і т. д.); 2) змінні, які вимірюють поведінку (участь у попередніх виборах і т. д.); 3) соціально-демографічні змінні.

Результати спроб побудови інструментів покращення передбачення явки в Україні

Під час планування панельного дослідження перед першим туром виборів Президента України 2004 р. ми розробили набір запитань, який вимірював змінні, що пов'язані з ймовірністю взяти участь у голосуванні. Респондентам були поставлені такі запитання: 1) про участь у загальнонаціональних виборах, які проходили в Україні; 2) про рівень зацікавленості в політиці; 3) про рівень упевненості респондента щодо свого наміру взяти участь у виборах; 4) про важливість результатів виборів для респондента; 5) про те, чи вважає респондент участь у виборах обов'язком кожного громадянина; 6) про те, чи буде подружжя респондента брати участь у виборах; 7) серія запитань про те, чи будуть голосувати чотири співрозмовники респондента. В індексі ймовірності проголосувати, побудованому на даних панельного дослідження, було подано такі альтернативи змінних (в дужках подано відсоток респондентів, які дали ствердну на запитання): високий рівень зацікавленості в політиці (16 %); згода з тим, що участь у виборах – обов'язок кожного громадянина (55 %); згода з тим, що результати виборів є дуже важливими (58 %); чи будуть усі чотири співрозмовники респондента голосувати (70 %); чи брав респондент участь в усіх виборах, в яких міг брати участь за віком (71 %); впевненість у своїй участі у виборах (76 %). Цей індекс складається з відповідей на шість запитань, відповідно може набирати шість значень (від 0 до 6). Поєднати ці відповіді в індекс можна двома способами: 1) побудувати адитивний індекс; 2) побудувати шкалу Гуттмана. Але, як показав наш аналіз, два способи побудови індексу дають дуже подібні ре-

зультати¹. Чи справді є залежність між тим, що респондент збирався взяти участь у виборах, але не проголосував, і значенням індексу (в таблиці наведено відповіді груп респондентів, виокремлених за однаковими значеннями індексу, на запитання про участь у виборах, яке ставилося в поствиборному дослідженні, проаналізовано відповіді тих респондентів, які в передвиборному дослідженні сказали, що будуть брати участь у виборах).

Таблиця 6. Відсоток тих, хто проголосував, у групах із різними значеннями індексу – дані поствиборного опитування

| Індекс ймовірності проголосувати | Кількість респондентів у групі | Відсоток | Так, я голосував 31 жовтня | Ні, я не голосував 31 жовтня |
|----------------------------------|--------------------------------|----------|----------------------------|------------------------------|
| 0 | 30 | 2% | 59% | 41% |
| 1 | 124 | 7% | 73% | 27% |
| 2 | 210 | 12% | 86% | 14% |
| 3 | 343 | 19% | 87% | 13% |
| 4 | 439 | 25% | 92% | 8% |
| 5 | 497 | 28% | 96% | 4% |
| 6 | 147 | 8% | 97% | 3% |
| Разом | 178 | 100% | 88% | 12% |

Дискримінаційна валідність побудованого індексу невисока – навіть у групах із найнижчим значенням індексу значна частина респондентів взяла участь, а групи з вищими значеннями індексу не дуже сильно відрізняються за відсотком тих, хто каже, що не голосував. Відповідно, не дуже обґрунтованим рішенням буде вилучення з аналізу навіть двох груп із найнижчими значеннями індексу – поміж членів цих груп більше половини виборців кажуть, що взяли участь у виборах. Розглянемо аналогічну таблицю з роботи [7].

Таблиця 7. Відсоток тих, хто проголосував у групах із різними значеннями індексу – дані поствиборного опитування

| Індекс ймовірності проголосувати | Відсоток | Респондент голосував | Респондент не голосував |
|----------------------------------|----------|----------------------|-------------------------|
| 1 | 11% | 4% | 96% |
| 2 | 2% | 13% | 87% |
| 3 | 10% | 9% | 91% |
| 4 | 4% | 30% | 70% |
| 5 | 4% | 28% | 72% |
| 6 | 9% | 58% | 42% |
| 7 | 10% | 62% | 38% |
| 8 | 51% | 75% | 25% |

Чому відрізняється якість дискримінації за індексом Траджетта-Такера та тим, що побудований нами? Ми не схильні вважати, що змінні,

¹ Коефіцієнт кореляції між двома індексами, побудованими двома способами – 0.86.

які використані в нашому індексі не пов'язані зі ймовірністю дотриматися наміру про участь у виборах. Але, у дослідженні Траджетта-Такера валідація факту голосування відбувалася через перевірку офіційної інформації про те, чи голосував конкретний виборець, а в нашому дослідженні висновки зроблено на основі даних поствиборного панельного дослідження. Причина низької якості дискримінації за допомогою нашого індексу полягає в низькій валідності даних, які отримуються в поствиборних панельних дослідженнях. Такі фактори загрожують валідності цих даних:

- 1) Зміщення, що викликані питаннями без відповідей.
- 2) Виборці схильні надавати неправдиву інформацію про свою участь у виборах – велика частина тих, хто насправді не голосував, кажуть, що зробили це.
- 3) Так званий ефект стимулу, який полягає в тому, що респонденти, які брали участь у передвиборному опитуванні з більшою ймовірністю беруть участь у виборах, ніж ті, кого не опитували.

Таким чином, побудований нами індекс є радше орієнтиром і засобом експериментування для дослідника, ніж точним інструментом. Наприклад, можна спробувати залучати до ймовірного електорату таку частку групи з певним значенням індексу, яка взяла участь у виборах згідно з результатами нашого аналізу (тобто, увести ваги). Для побудови надійнішого показника потрібно мати можливість отримати надійні дані з офіційних джерел про те, які респонденти взяли участь у виборах, а які ні.

Ми також побудували на даних панельного дослідження рівняння логістичної регресії, де залежною змінною було те, чи дотримався респондент наміру взяти участь у виборах, чи ні, а незалежними – змінні, які використовувалися під час побудови індексу, та демографічні змінні. Модель виявилася такою:

Таблиця 8. Модель логістичної регресії для розрахунку ймовірності взяти участь у виборах.

| | B | S.E. | Wald | Sig. | Exp(B) |
|--|--------------|--------------|--------------|--------------|---------------|
| Bik | 0,010 | 0,005 | 3,080 | 0,079 | 1,010 |
| Участь у всіх попередніх виборах | -0,656 | 0,204 | 10,306 | 0,001 | 0,519 |
| Кількість співрозмовників, які беруть участь у виборах | -0,396 | 0,191 | 4,312 | 0,038 | 0,673 |
| Упевненість у своїй участі у виборах | 0,737 | 0,208 | 12,514 | 0,000 | 2,089 |
| Висока важливість виборів | -0,280 | 0,124 | 5,090 | 0,024 | 0,756 |
| Константа | -0,874 | 1,090 | 0,642 | 0,423 | 0,417 |

Попри те, що модель виявилася статистично значущою¹, вона пояснює невелику частку дисперсії² та не є надто корисною для прогнозування – тільки 12 % респондентів, які не взяли участі у виборах, задекларували такий намір у передвиборному дослідженні, були правильно класифіковані моделлю. Зважування на ймовірність взяти участь у виборах практично не змінило рейтинги кандидатів у президенти. Тобто, побудована нами модель має низьку прогностичну спроможність. Ми вважаємо, що основною причиною, як і у випадку з індексом, є те, що ми користувалися даними поствиборного панельного дослідження. Для того, щоби перевірити цю гіпотезу, ми на даних National Election Study побудували модель, повністю ідентичну моделі Петровича [9]. Але, якщо Петрович валідизував свою модель за допомогою даних офіційної статистики, то ми користувалися тільки відповідями респондентів у поствиборному дослідженні. Виявилось, що модель, яка, за свідченнями Петровича, достатньо ефективно передбачає поведінку виборців (участь або неучасть у виборах) у разі, коли вона побудована з використанням надійної інформації (з офіційних джерел) про залежну змінну (голосував респондент чи ні), неефективно працює, якщо її будувати на даних поствиборного опитування. Як і у випадку українського зразка, модель Петровича, побудована на панельних даних, правильно визначає як «не виборців» тільки 12 % тих, хто справді не голосував, хоча за умови використання офіційної інформації працює набагато краще.

Ми можемо зробити висновок, що передбачення того, хто саме з респондентів візьме участь у виборах, є дуже важливим завданням для соціологів, і для його розв'язання існують достатньо перспективні методи. Головна проблема – знайти розв'язання того, що респонденти надають неправильну інформацію про свої наміри голосувати. Для цього дослідникам обов'язково треба залучити до опитувальників передвиборних досліджень набір запитань, які вимірюють змінні, пов'язані з ймовірністю проголосувати. За допомогою цієї інформації можуть бути обчислені ймовірності респондентів справді взяти участь у виборах.

Проте ключовим для успішної побудови потрібних рівнянь та індексів є доступ дослідників до валідної інформації про голосування респондентів – на жаль, дані панельних досліджень є недостатньо надійними.

¹ Sig = .000

² Nagelkerke R Square дорівнює .072

1. Crespi I. Sources of accuracy and errors in Pre-Election Polling. N. Y.: Sage. 1981.
2. Freedman and Goldstein. Building a probable electorate from Preelection polls. // Public Opinion Quarterly Volume 60: pp. 574–587.
3. Bernstein R., Chadcha A., Montjoy R. Overreporting Voting: Why It Happens and Why It Matters. //Public Opinion Quarterly Volume 65: pp. 22–44.
4. Lau R. An analysis of the Accuracy of «Trial Heat» Polls During the 1992 Presidential election. // Public Opinion Quarterly Volume 58: pp. 2–20.
5. Вибори 2002 : передвиборні рейтинги і результати голосування. <http://www.kiis.com.ua/txt/pr.htm>
6. Ulises Beltrán and Marcos Valdivia. Accuracy and Error in Electoral Forecasts: the Case of Mexico. International Journal of Public Opinion Research., SUMMER 1999; 11: pp. 115–134.
7. Traugott M. and Tucker C. Strategies for Predicting whether a Citizen will Vote and Estimation of Electoral Outcomes. // Public Opinion Quarterly Volume 48: pp 330–348
8. Perry P. Certain Problem in Election Survey Methodology. // Public Opinion Quarterly Volume 1979, Issue 5 pp. 312–325.
9. Petrocik R. J. An algorithm for Estimating Turnout as a Guide to Predicting Elections. // Public Opinion Quarterly Volume 55, pp. 643–647.
10. Murray R., Riley K., Scime A. Pre-election polling : identifying likely voters using iterative expert data mining // Public Opinion Quarterly Volume 2009, 73, Number 1, 2 April 2009 , pp. 159–171.

Bolshov E. S.

FORECASTING TURNOUT AS A FACTOR OF ACCURACY OF ELECTION RESULT'S FORECASTING

Impact of turnout forecasting accuracy on accuracy of election results's forecasting is analyzed in article. Methods of detecting «likely voters» are studied. Variables that can be used for these purposes are examined. Results of methods testing are described.

Keywords: election result forecasting, turnout forecasting, «likely voters»

УДК 316.74:2

Богданова О. В.

НАСКІЛЬКИ РЕЛІГІЙНИМ Є УКРАЇНСЬКЕ СУСПІЛЬСТВО? ПОРІВНЯЛЬНИЙ АНАЛІЗ ПОКАЗНИКІВ РЕЛІГІЙНОСТІ ЗА ДАНИМИ МІЖНАРОДНОЇ ПРОГРАМИ СОЦІАЛЬНИХ ОПИТУВАНЬ (ISSP, 2008)

У статті запропоновано оцінку рівня релігійності українського суспільства за результатами опитування в рамках Міжнародної програми соціальних опитувань (International Social Survey Programme, ISSP) 2008 р., реалізованої в Україні Інститутом соціології НАН України та Київським міжнародним інститутом соціології. Оцінку здійснено за такими показниками релігійності, як віра в продовження свідомого існування після смерті, віра в Бога та релігійна самоідентифікація; проаналізовано валідність відповідних показників релігійності.

Ключові слова: релігія, релігійність, вимірювання релігійності, оцінка релігійності, Україна.

У науковій літературі розрізняють низку параметрів, за якими можна оцінювати рівень релігійності: самооцінка рівня релігійності, самооцінка важливості релігії, самоідентифікація з певною релігією, релігійні вірування (віра в Бога або богів, у продовження свідомого існу-

вання після смерті, специфічні догмати), дотримання релігійних практик (поміж яких індивідуальна та колективна молитви, специфічні ритуали). Утім, у жодному опитуванні ми не можемо повноцінно розкрити кожен із цих параметрів; відповідно, перед дослідниками незмінно постає